Playing TETris with DNA modifications

Benjamin Delatte, Rachel Deplus & François Fuks*

Abstract

Methylation of the fifth carbon of cytosine was the first epigenetic modification to be discovered in DNA. Recently, three new DNA modifications have come to light: hydroxymethylcytosine, formylcytosine, and carboxylcytosine, all generated by oxidation of methylcytosine by Ten Eleven Translocation (TET) enzymes. These modifications can initiate full DNA demethylation, but they are also likely to participate, like methylcytosine, in epigenetic signaling per se. A scenario is emerging in which coordinated regulation at multiple levels governs the participation of TETs in a wide range of physiological functions, sometimes via a mechanism unrelated to their enzymatic activity. Although still under construction, a sophisticated picture is rapidly forming where, according to the function to be performed, TETs ensure epigenetic marking to create specific landscapes, and whose improper build-up can lead to diseases such as cancer and neurodegenerative disorders.

Keywords DNA modifications; epigenetics; human diseases; hydroxymethylcytosine; TET proteins

DOI 10.15252/embj.201488290 | Received 20 February 2014 | Revised 20 March 2014 | Accepted 8 April 2014 | Published online 13 May 2014

The EMBO Journal (2014) 33: 1198–1211

See the Glossary for abbreviations used in this article.

Introduction

DNA methylation was observed more than a century ago by W. G. Ruppel, who discovered in Mycobacterium tuberculosis, the causative agent of tuberculosis, a non-canonical nucleotide that he believed to be a methylated pyrimidine (Ruppel, 1898). Since this initial discovery, the modified base has been identified as 5-methylcytosine (5mC), and a plethora of papers have shown it to play many roles in both normal physiology and disease. This epigenetic cytosine (5mC), and a plethora of papers have shown it to play initial discovery, the modified base has been identified as 5-methylcytosine (Ruppel, 1898). Since this initial deamination of 5hmC to 5-hydroxymethyluridine (5-hmU) by TDG glycosylase, followed by DNA repair. Alternatively, an initial demethylation of 5hmC to 5-hydroxymethyluridylate (5-hmU) by AID/APOBEC enzymes followed by BER also leads to complete DNA demethylation (Guo et al., 2011). Intriguingly, decarboxylation of 5caC has been observed in mouse embryonic stem cell (mESC) extracts, but no putative decarboxylase has yet been discovered (Schiesser et al., 2012). In 2009, two seminal papers reported the existence of another modified form of cytosine: 5-hydroxymethylcytosine (5hmC). This modified base is generated by the Ten Eleven Translocation (TET) enzymes through oxidation of 5mC (Kriaucionis & Heintz, 2009; Tahiliani et al., 2009). The TET enzyme family comprises three TETs (TET1, TET2, TET3), which all derive from a common ancestor that was triplicated in a jawed vertebrate (Pastor et al., 2013). They use Fe²⁺ and α-ketoglutarate for catalysis, and all three TETs can also iteratively oxidize 5-hydroxymethylcytosine to 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Tahiliani et al., 2009; Ito et al., 2011; Blaschke et al., 2013; Minor et al., 2013; Yin et al., 2013). In tissues, the abundance of these DNA modifications tends to vary in the order 5mC > 5hmC > 5fC > 5caC, where both 5fC and 5caC abundance are respectively found at nearly one and two orders of magnitude lower than 5hmC. In two recent studies, TET-mediated catalysis was further investigated by analysing the crystal structures of human TET2 and an amoebal homologue of mammalian TET1 (NgTET1 from Naegleria gruberi). It seems that the negatively charged DNA binds the basic amino acids located at the surface of the TET protein and that the enzyme flips the methylcytosine out of the DNA double helix into its double-stranded β-helix catalytic pocket where it is stabilized by hydrogen bonds. DNMTs also show such a base-flipping mechanism. Importantly, the cysteine-rich DNA binding domain located in the carboxyterminal portion of human TET2 also plays a role in stabilizing the double-stranded β-helix domain (Hu et al., 2013; Hashimoto et al., 2014). Through their catalytic activity, the TETs act as initiators of DNA demethylation, either via passive dilution of 5hmC with DNA replication as seen in maturating primordial germ cells (PGCs) and fertilized zygotes, or through the action of DNA repair enzymes (the base excision repair (BER) machinery), which enzymatically remove the modified base and replace it with a cytosine (Guo et al., 2011; He et al., 2011; Maiti & Drohat, 2011; Zhang et al., 2012). This last process can occur through direct removal of 5fC and 5caC by the TDG glycosylase, followed by DNA repair. Alternatively, an initial deamination of 5hmC to 5-hydroxymethyluridylate (5-hmU) by AID/APOBEC enzymes followed by BER also leads to complete DNA demethylation (Guo et al., 2011). Intriguingly, decarboxylation of 5caC has been observed in mouse embryonic stem cell (mESC) extracts, but no putative decarboxylase has yet been discovered (Schiesser et al., 2012). Lastly, Chen et al (2012) reported in 2012 that DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases, and that incubation of these methyltransferases with 5hmC-containing PCR products leads to partial...
dehydroxymethylation, this activity being increased in an oxidizing environment.

In summary, many routes of TET-triggered DNA demethylation (replication dependent, BER, deamination followed by BER, decarboxylation or dehydroxymethylation) are opened by a first oxidation of 5mC to 5hmC. Yet we should bear in mind that gene-rich regions in some tissues, such as those of the nervous system, display significant “stable” enrichment in hydroxymethylcytosine (Kriaucionis & Heintz, 2009), suggesting that TET-mediated 5hmC formation not only primes demethylation but also creates a unique epigenetic landscape in these cells.

Previous reviews discussing DNA modifications created by TET proteins have notably dealt with general aspects, their genomic distribution, and their roles in embryonic development (e.g. Delatte & Fuchs, 2013; Pastor et al, 2013; Wu & Zhang, 2014). The present review covers more recent discoveries in the exciting world of TETs. It first offers a glimpse at new TETs modes of action that are now surfacing thanks to studies providing novel insights into TET regulation, TET partners, and readers of DNA modifications. The focus then shifts to the roles played by TETs in key physiological processes and to the involvement of TET malfunction in disease, with special emphasis on non-cancer pathologies.

Novel insights into the mechanisms of TET-mediated DNA modifications

While our picture of how TET enzymes contribute to the dynamic methylation/demethylation cycle is still blurred, researchers have begun to lift the veil on how TET enzymatic activity is regulated, how TETs are targeted to precise DNA sequences, and how the 5mC oxidation signal is further read by “interpreters”. Below, we describe significant recent advances in this area.

Modes of direct regulation of TET enzymes

Direct regulation of TET enzymatic activity should be a rather straightforward means of modulating DNA modification levels (Fig 1). For 5mC oxidation, the TET proteins use Fe(II) and O2 as cofactors along with α-ketoglutarate, a metabolite of the Krebs cycle, which is converted to succinate (Tahiliani et al, 2009). In several cancers, including 80–90% of adult grade II/III gliomas (Losman & Kaelin, 2013), the isocitrate dehydrogenases 1 and 2 (IDH1 and IDH2) are mutated and the oncometabolite 2-hydroxylglutarate (2-HG) is produced. As 2-HG can compete directly with α-ketoglutarate, it unsurprisingly inhibits various dioxygenases, including the TET enzymes (Xu et al, 2011a). As the TET and IDH proteins act in concert, acute myeloid leukaemias with IDH mutations display both a global and a specific hypermethylation signature, possibly due to decreased TET activity (Figueroa et al, 2010).

Novel insights into the direct regulation of TET activity were obtained in 2013. Several groups independently reported that ascorbic acid (vitamin C) can significantly enhance TET activity in vitro, in cultured cells, and in vivo (Blaschke et al, 2013; Minor et al, 2013; Yin et al, 2013). Vitamin C deficiency has long been known to result in scurvy, a human disease caused by impaired collagen synthesis. The characteristic symptoms of scurvy are due to decreased prolyl 4-hydroxylase activity and incomplete hydroxylation of collagen fibres (Choikier et al, 1983; Peterkofsky, 1991). In mouse embryonic fibroblasts (MEFs), Minor et al (2013) observed a 50% increase of the 5hmC level after 1 h of treatment with 10 μM ascorbate, and up to a five-fold increase after longer exposure times. Yin et al further demonstrated both in cultured cells and with purified TET domains that vitamin C causes rapid formation of 5hmC, 5fC, and 5caC, probably by reducing Fe(III) back to Fe(II) after enzymatic catalysis. In keeping with the ability of the 5hmC-5fC-5caC pathway to induce demethylation, mESCs treated with vitamin C display a significant concomitant demethylation. These results were confirmed in vivo in mice deficient in ascorbic acid synthesis.
(Yin et al., 2013), and genome-wide mapping of 5hmC after vitamin C treatment of mESCs evidenced a global increase in hydroxymethylation and a decrease in 5mC (Blaschke et al., 2013; Yin et al., 2013). Interestingly, Blaschke et al. found vitamin C to both induce DNA demethylation and increase expression of germline genes, whereas some regions such as imprinted loci and intracisternal A particle (IAP) retroelements proved refractory to DNA demethylation. Furthermore, although various chemical reducing agents (GSH, vitamin B1, vitamin E) appeared not to affect 5-methylcytosine oxidation patterns, L-cysteine was shown to increase 5hmC by nearly 20% and 5FC by 50% (Blaschke et al., 2013; Yin et al., 2013). Similarly, hydroquinone, a potent reducing agent used for skin depigmentation, significantly increases hydroxymethylcytosine (Coulter et al., 2013). It remains to be seen whether oxidizing agents negatively impact TET activity, perhaps by oxidizing the Fe(II) catalytic centre. Unpublished data from our lab point to inactivation of the TET enzymes by hydrogen peroxide treatment, and we have observed a significant decrease in 5hmC in both cells treated with various oxidants and Gpx1/-/- knockout mice (Gpx enzymes being the major natural antioxidants in the cells). As oxidative stress is involved in various diseases such as cancers and neuronal disorders, it will be important to assess whether alteration of global and local 5hmC patterns might explain the manifestations of these life-threatening diseases (B. Delatte & F. Fuks, unpublished data).

Besides regulation by small molecules, recent reports indicate that TET proteins can be post-transcriptionally regulated by microRNAs (miRNAs or miRs) (Cheng et al., 2013; Fu et al., 2013; Morita et al., 2013; Song et al., 2013a,b). For example, Morita et al. (2013) have shown that miR-29-family of microRNAs can repress TET1 and thymine-DNA glycosylase (T DG) without greatly altering DNA methylation, possibly because they can also downregulate DNMT3A and DNMT3B. Zhang et al. (2013a) have found miR-29 to inhibit TET1-TET2 and TET3 as well as T DG, this being accompanied by a small but significant decrease in hydroxymethylcytosine. MiR-26 can also repress all three TETs and T DG, emphasizing a quite complex regulation of DNA demethylation enzymes by small RNAs (Fu et al., 2013). Interestingly, transgenic mice overexpressing miR-26 display a decreased level of 5hmC and an increased proliferation and differentiation of pancreatic cells in vivo. This suggests that TETs might play a role in pancreatic differentiation and also in pancreas-related diseases such as cancer and diabetes (Fu et al., 2013). As miR-26 and many other miRs are deregulated in human malignancies and can function either as oncogenes or as tumour-suppressor genes (Fu et al., 2013), and as a decreased hydroxymethylcytosine level is suggested to be a hallmark of cancer (see below; cf. also Delatte & Fuks, 2013), TET proteins might be directly downregulated by miR networks during tumorigenesis. It is noteworthy in this regard that overexpression of miR-22 in the haematopoietic or mammary gland compartment inhibits TET-mediated hydroxymethylation and promotes leukaemogenesis or breast cancer, respectively. In fact, myelodysplastic patients displaying a decreased level of TET2 show poor survival and an inverse correlation between the TET2 and miR-22 levels. A similar anti-correlation has been observed in breast cancer patients exhibiting reduced TET1/TET2 expression (Song et al., 2013a,b). TET2 being a target of miR-29 and miR-26 has been recently confirmed by Chen et al., who studied the effects of overexpressing nearly 500 miR constructs in H293T cells. They observed no deregulation by miR-22, possibly for technical reasons such as the length of the 3′-untranslated regions used in their experiments. Remarkably, they identified another ten miR families that regulate TET2, notably miR-101 and miR-125, which induce abnormal haematopoiesis and a malignant phenotype in xenograft experiments, directly via TET2 downregulation (Cheng et al., 2013). These reports highlight the possible role of miRNAs in mediating the post-transcriptional regulation of TETs during pathogenesis.

Another mode of TET regulation to be considered is post-translational modifications (PTM). Covalent PTMs (such as acetylation, methylation, phosphorylation, and glycosylation) are well known to provide a vast indexing potential and to expand the range of functions of histones and non-histone proteins. Little has been published on the possible regulation of TETs by post-translational modifications. We do know that all three TETs can be modified by the O-linked N-acetylg glucosamine (O-GlcNAc) transferase (OGT) (Shi et al., 2013; Vella et al., 2013; Ito et al., 2014). This glycosyltransferase adds O-GlcNAc moieties on various cytosolic and nuclear proteins (Hu et al., 2010). Protein O-glycosylation is known to antagonize phosphorylation, by direct competition for site-occupancy at a single hydroxyl group or via steric hindrance at proximal sites in the same polypeptide (Wang et al., 2008). As glycosylation of threonine 535 of TET1 enhances the protein’s stability (Shi et al., 2013), it is tempting to propose that phosphorylation of this residue in the absence of GlcNAc might flag the protein for ubiquitin-mediated degradation and thereby cause a significant decrease in 5mC oxidation marks (Shi et al., 2013). One might expect TETs, like many other epigenetic enzymes (Denis et al., 2011), to be decorated by various PTMs affecting their enzymatic activity.

In summary, the body of available evidence suggests that small molecules, microRNAs, and PTMs may regulate hydroxymethylcytosine levels by directly modulating TET expression and/or enzymatic activity. Future studies may bring to light additional modes of regulation, notably involving direct amino acid modifications.

An expanding TET interaction network

Since the original discovery of TETs in 2009, another approach to elucidating TET functions has been to identify proteins with which these enzymes interact. The search for TET interactants has brought to light a plethora of transcription-related factors: transcription factors/nuclear receptors (CXXC4, NANO G, PPARY, PU.1, EBF1, PRDM14) and chromatin-associated proteins involved in transcriptional activation (OGT and SET1/COMPASS complex) or repression (SIN3A/HDACs, NURD). Below we describe in more detail what we know about these TET partners and how they might help TETs to perform their functions (Fig 2).

The strongest TET partner identified so far is OGT, which we and others have recently reported (Chen et al., 2012; Deplus et al., 2013; Shi et al., 2013; Vella et al., 2013; Ito et al., 2014). Recent evidence of an epigenetic function for OGT has emerged with the discovery that it can glycosylate histone H2B on serine 112, causing its ubiquitylation and the subsequent trimethylation of histone H3 at lysine 4 (H3K4me3). This modification is known to “decorate” active genes and to be mainly deposited by the SET1/COMPASS complex (Fujiki et al., 2011, 2013). Depending on TETs and OGT abundance and on the cell context, TET1, TET2, and TET3 all seem able to interact
physically with OGT. The published data now enable us to propose a “hydroxymethylation independent” model of gene activation, in which TETs associate with OGT on promoters (mainly at CpG rich islands) and enhance its glycosyltransferase activity. This would in turn favour glycosylation of serine 112 of H2B (H2BS112GlcNAc) and ubiquitinylation of lysine 120 (H2BK120ub). In parallel, TET-promoted OGT activity would stabilize by glycosylation the SET1/COMPASS complex, this resulting in SET1/COMPASS binding to H2BK120ub, trimethylation of H3K4, and transcriptional activation (Chen et al., 2013; Deplus et al., 2013; Vella et al., 2013). Efforts should be made to unravel the physiological role(s) of this strong TET-OGT connection. Studies on neuronal cells might prove particularly fruitful, as both OGT and the TETs appear most abundant in these cells and are essential to neuronal development (Okuyama & Marshall, 2003; Kriaucionis & Heintz, 2009).

Genome-wide studies have shown that TET1 is associated not only with active “H3K4me3 only” promoters but also with bivalent poised “H3K4me3/H3K27me3” promoters and repressed “H3K27me3 only” promoters. In mESCs, furthermore, TET1 knockdown leads to activation of a number of genes, and this suggests that TET1 may have a transcriptional repressor role (Wu et al., 2011). In 2011, Wu et al. found TET1 depletion to decrease binding of the H3K27me3 responsible enzyme EZH2 to co-bound genes. While TET1 has been reported to interact with EZH2 (Cartron et al., 2013), other investigators have failed to observe such an association (Williams et al., 2011a; Wu et al., 2011; our unpublished work). It thus seems likely that the TET1-EZH2 link is an indirect one, in which TET1 would “prepare the ground” for EZH2 binding by “shaping” the promoter’s DNA modification state. In 2011, the Helin’s laboratory described the SIN3A corepressor complex as a binding partner of TET1 (Williams et al., 2011b). SIN3A is known to associate with histone deacetylases HDAC1 and HDAC2, enzymes that maintain histones in a deacetylated form, thereby repressing transcription. Williams et al. (2011b) further showed that TET1 clusters around transcription start sites with SIN3A, recruits SIN3A to a subset of promoters and induces transcriptional silencing of the co-bound genes. HDAC 1 and 2 are parts of other corepressor complexes such as NuRD or coREST, and NuRD has also been found to interact with TET1 in mESCs (Yildirim et al., 2011; Kelly & Cowley, 2013; Shi et al., 2013). As described below, the MBD3 NurD component can bind 5-hmC via its MBD domain and the genome-wide profile of MBD3 is similar to that of TET1 (Yildirim et al., 2011). These combined results thus constitute evidence of a link between TET1, 5mC, and histone deacetylase systems in transcriptional gene repression. TETs are
likely to be found to cooperate with yet other chromatin complexes - both activating and repressive - in controlling gene expression. TET1 and TET3 can bind DNA via their N-terminal CXXC and C-terminal cys-rich domains, but TET2 has lost its CXXC domain during evolution, which is now encoded by a separate gene called C-terminal cys-rich domains, but TET2 has lost its CXXC domain likely to be found to cooperate with yet other chromatin complexes.

New insights into TET function

During differentiation, TET1/TET2 proteins have been shown to facilitate reprogramming of neural stem cells in a manner that depends on methylcytosine hydroxylase activity (Costa et al., 2013). TET recruitment has also been observed during differentiation. One example concerns the differentiation of pre-adipocytes to adipocytes, triggered by the nuclear receptor PPAR. During adipogenesis and in a ligand-dependent manner, PPARγ switches binding partners from a corepressor to a co-activator complex on the PPAR-response elements (PPREs) of numerous genes expressed specifically in adipocytes. PPARγ poly(ADP-ribosyl)ation (PARylation), a post-translational protein modification mediated by PARPs, is known to regulate PPARγ transcriptional activity, and this may involve DNA demethylation (Fujiki et al., 2013). In a recent study, Fujiki et al showed that during differentiation, TET1 interacts with PPARγ after PARylation of the nuclear receptor. They further revealed that TET1 and TET2 can bind to PAR polymers and propose a model in which active DNA demethylation of key adipocyte-specific genes occurs via direct binding of TET1 and TET2 to the modified PPARγ receptor (Fujiki et al., 2013). During osteoclast differentiation, likewise, the transcription factor PU.1 interacts with TET2 and recruits it to genes that end up hypomethylated (de la Rica et al., 2013). Interestingly, PU.1 is required to generate most myeloid lineages (including macrophages, neutrophils, and dendritic cells) and for B-cell differentiation; it can be suppressed or mutated in leukaemia (Shereen & Rodney, 2011). It can also interact with the spliceosome machinery and participate in alternative splice site selection (Guillouf et al., 2006). These observations suggest an unanticipated involvement of the TET2-PU.1 association in cancer and in essential processes such as mRNA metabolism. On the basis of a search for consensus binding sites surrounding hypermethylated positions in a meta-analysis of IDH-mutant cancers, Guilhamon et al identified EBF1 as an interacting partner of TET2 in a human chondrosarcoma cell line. Further experiments showed that EBF1 and TET2 co-occupy certain genomic loci (Guilhamon et al., 2013). It is known that EBF1 is required for B lymphopoiesis and a significant fraction of genes, co-bound by EBF1 and the transcription factor E2-alpha (E2A), correlates with histones modifications of activated chromatin but also with poised or repressed regions in pro-B cells (Hagman et al., 2012). Knowing the interaction with EBF1 and that TET2 has a critical function in the haematopoietic compartment, it will be of great interest to assess its role in B-lymphocytes signaling and differentiation (Delatte & Fuks, 2013).

TET1 and TET2 also interact with the transcriptional regulator PRDM14, whose overexpression in mESCs leads to increased hydroxymethylation, demethylation, and subsequent activation of a set of pluripotency, germine, and imprinted genes (Okashita et al., 2013). As mentioned above, PGCs undergo global demethylation during their maturation, as a result of global conversion of methylcytosine to hydroxymethylcytosine and replication-dependent dilution. During this global wave of methylation resetting, imprinted loci also become demethylated, and TET1 has been found to play a role in erasing paternal and maternal genomic imprints, this being particularly evident in a TET1 knockout mouse which also display a defect in meiosis of developed gametes and some developmental arrest between E16.5 and E18.5. As PRDM14 is highly expressed in PGCs and interacts with TET1, it is tempting to speculate that these two proteins might participate together in imprinting erasure during PGC reprogramming (Okashita et al., 2013; Yamaguchi et al., 2013).

All in all, the functional diversity of the TET interaction network (Fig 2) suggests that the TET proteins are more versatile than initially believed and participate in other processes besides demethylation of CpG dinucleotides. ChiP-seq experiments on histone marks in relevant tissues will help to unravel the physiological roles of the TETs beyond differentiation and reprogramming. Also, metabolic and signalling pathways (such as those involved in OGT metabolism or Wnt/β-catenin) are likely to modify the TET-related epigenome. An important future task could thus be to compare maps of OGT or CXXC4 with maps of TET2 and 5hmC, for example, under various conditions affecting these pathways.

Readers of oxidized methylcytosines

In mammals, 5hmC levels vary according to the tissue and cell type. The brain and spinal cord show the highest levels, and organs such as the testes, spleen, and thymus display the lowest. An intermediate level is recorded in embryonic stem cells were approximately 0.1% of all bases in these cells are hydroxymethylcytosines (Globisch et al., 2010; Ito et al., 2010; Szewczuk et al., 2010). The abundance of this oxidized methylcytosine, particularly in the brain, has led the scientific community to seek proteins that can bind to 5hmC and interpret/translate the encoded information. Several studies have shown that methyl-binding proteins (MBDs) can to some extent bind to 5hmC DNA, most often within a hemi-hydroxymethylated template and with lower affinity than to fully methylated DNA. This is the case of MeCP2, MBD1, MBD2, and MBD4 (Jin et al., 2010; Mellen et al., 2012; Otani et al., 2013; Hashimoto et al., 2014). MBD3 has been known for some time not to bind 5mC on its own, but in 2011, Yildirim and coworkers reported that the repressive NurD complex can bind both 5hmC-containing and non-modified DNA through its MBD3 subunit (Yildirim et al., 2011). The jury is still out as to whether MBD3 is a true hmC reader, as it has not been detected by others (Spruijt et al., 2013). The fact that MeCP2 can recognize 5hmC in vitro will force the community to re-evaluate the MethylCap assay, which employs the MeCP2 MBD domain to interrogate 5mC in genomic DNA. This applies particularly to tissues such as brain tissue, where almost half of the methylcytosine is in fact hydroxymethylcytosine (Brinkman et al., 2010). Further in silico and in vitro studies have examined whether the DNMT1 partner UHRF1 might also bind to hemi- or fully hydroxymethylated DNA via its SRA domain. It appears that UHRF1 does bind 5hmC-containing DNA, but the function of this binding is not yet clear. The protein is known to flip the base out of the DNA double helix. This might enhance further oxidation of 5hmC to 5fC and 5caC by
making the hydroxymethylated base more accessible to the TET enzymes (Frauer et al., 2011). Accordingly, H293T cells overexpressing the UHRF1 homologue UHRF2, which also binds hydroxymethylcytosine, display oxidation of 5mC to 5hmC, with a remarkable increase in 5fC and 5caC (Spruijt et al., 2013).

The above-mentioned studies mostly describe in vitro approaches where recombinant proteins are incubated with modified DNA fragments and then used in electrophoretic mobility shift assays. Two groups have reported identifying readers of 5hmC, 5fC, and 5caC by quantitative mass spectrometry (Iurlaro et al., 2013; Spruijt et al., 2013). Interestingly, the first group found some readers to be specific to mESCs, neural progenitor cells (NPCs), or adult mouse brain tissue, while the second revealed proteins whose binding to 5hmC or 5fC depends on the CpG composition. This indicates (i) that hydroxymethylcytosine probably behaves differently in different cell types and (ii) that some presumptive readers bind oxidized methylcytosines in a sequence-specific manner. The screens also identified several specific 5fC and 5caC readers, among them TDG, supporting the view that this protein can act as a DNA demethylase (Iurlaro et al., 2013; Spruijt et al., 2013).

In summary, specific proteins are recruited to 5hmC, 5fC, or 5caC sites, and some of them are likely to regulate gene expression or to drive active DNA demethylation. More in-depth characterization by ChIP-seq and the use of knockout mice should help to show whether the proteins identified so far are true “readers” that can interpret 5hmC, 5fC, or 5caC modifications and translate them into an activating or repressing signal, or merely “binders” with some affinity for modified cytosines.

**Emerging roles of TETs in normal physiology and disease**

**TETs/hmC in cell reprogramming and differentiation**

The TET proteins participate in two waves of global DNA demethylation during reprogramming of primordial germ cells as well as of paternal pronucleus in the zygote. Such reprogrammings are needed to ensure resetting of the epigenomes for pluripotency (Gu et al., 2011; Inoue et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011; Hackett et al., 2012; Nakamura et al., 2012). When PGCs migrate to the future gonads, high levels of TET1 and TET2 allows conversion of 5mC into 5hmC and this, not only globally but also more locally (e.g. at genomic imprinted loci). This is followed by subsequent passive “replication-coupled” dilution of 5hmC during several round of cell divisions with a progressive loss of hydroxymethylation (Hackett et al., 2012). A second wave of global DNA demethylation appears when the spermatoozon fertilizes the oocyte. In this case, TET3 hydroxylates the paternal DNA while PGC7 (also called DPPA3/STELLA) maintains normal methylation levels on the maternal genome as well as on imprinted genes by binding to the repressive H3K9me2 histone mark and by repelling TET3 (Szabo & Pfeifer, 2012). Worth mentioning, it has been observed in the paternal pronucleus, 5hmC is further oxidized into 5fC and 5caC and that those modifications are also passively diluted through cell divisions (Gu et al., 2011; Inoue & Zhang, 2011; Inoue et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011; Nakamura et al., 2012).

The initial discovery of TET enzymes in ESCs has led many groups to assess their importance in pluripotency maintenance and cell differentiation. Although some reports suggest a role of TET1 in maintenance of pluripotency, especially via direct binding and demethylation of the NANOG promoter, other papers do not support this view (Ito et al., 2010; Ficz et al., 2011; Koh et al., 2011; Williams et al., 2011a; Wu et al., 2011; Xu et al., 2011b; Doege et al., 2012; Freudenberg et al., 2012). Furthermore, in vivo results obtained with TET1 knockout and TET1/TET2 double-knockout mice (under conditions where TET3 is unlikely to compensate significantly for the knockout) suggest that ESCs stemness can be maintained without these methylcytoxigenases (Dawlaty et al., 2011, 2013). In fact, although a fraction of the double KO mice displayed midgestation abnormalities and perinatal lethality, viable and fertile mice were also obtained. Those however display a reduced size and weight as compared to their wt counterparts and the females also show smaller ovaries and reduced fertility. This indicates an overtly normal development with variable phenotypes probably due to variable penetrance of epigenetic defects. This double knockout is in sharp contrast with the TET3 knockout embryos from Xu’s group that show apparent lethality either around embryonic day 11.5 or after birth, highlighting the crucial role of TET3 in reprogramming the paternal pronucleus after fertilization (Gu et al., 2011). While the jury is still out on whether TETs play a role in maintaining pluripotency in ESCs, another phenomenon involving TETs is coming to light: reprogramming of developmentally committed cells such as MEFs to produce induced pluripotent stem cells (iPSCs). iPSCs, phenotypically comparable to ESCs, can be produced in vitro upon overexpression of the Yamanaka “OSKM” factors (OCT4, SOX2, KLF4 and C-MYC). During this process profound epigenetic reprogramming takes place, with loss of methylation on the promoters of several pluripotency genes (Takahashi & Yamanaka, 2006). Several studies have stressed the importance of TETs in iPSC reprogramming. In the first of these studies, TET2 and PARP1 were identified as key factors in this process: both are upregulated in MEFs upon OSKM treatment, and both bind to ESRRB and NANOG to shape the chromatin landscape essential for correct gene expression and iPSC reprogramming (Doege et al., 2012). As mentioned earlier, another report revealed that NANOG can interact with TET1 and TET2 and facilitate reprogramming of neural stem cells in a TET-activity-dependent manner (Costa et al., 2013). Remarkably, Gao et al. (2013) were able to induce iPSC formation in experiments where they replaced OSKM with “TSKM” (TET1, SOX2, KLF4, and C-MYC), probably because OCT4 is a direct binding target of TET1. These results highlight the importance of TET1 and TET2 in the demethylation and activation of key pluripotency genes such as ESRRB, NANOG, and OCT4 during reprogramming of differentiated cells, and thus suggest an important role for methylcytoxigenases in regenerative medicine studies.

Recent reports pinpoint a participation of TETs in terminal differentiation as well. In 2013, Hahn et al found the 5hmC pattern to be very dynamic during NPC differentiation to neurons. In gene bodies they observed a global increase in 5hmC without any subsequent change in DNA methylation. They further showed that TET2 and TET3 are required for correct NPC differentiation and that their loss leads to abnormal accumulation of cell clusters along the radial axis in the intermediate and ventricular zones (Hahn et al., 2013). A similar increase in hydroxymethylcytosine is seen during olfactory neuron differentiation, and TET3 overexpression disrupts both the olfactory receptor expression pattern and the targeting of axons to the olfactory bulb (Colquitt et al, 2013). These two studies strongly...
suggest that the hydroxymethylation landscape is important for the correct formation of brain tissues, as already reported by Shi’s group, which found TET3 to be required for normal development of the eye and brain in *Xenopus laevis* (Xu *et al.*, 2012). Three recent papers have shed further light on the involvement of TET enzymes in brain development (Kaas *et al.*, 2013; Rudenko *et al.*, 2013; Zhang *et al.*, 2013b). Zhang *et al.* found that TET1 is needed for NPC proliferation and that in vivo depletion of TET1 affects neuron production in the hippocampus. Further characterization of the mice in Morris water maze experiments revealed that their short-term memory was reduced (Zhang *et al.*, 2013b). Rudenko *et al.* (2013) reached different conclusions: in fear conditioning and Morris water maze tests, they found a normal neuronal density in the hippocampus of TET1-knockout mice, with no short-term memory impairment but with a significant decrease in memory extinction (which affects long-term memory), possibly due to alteration of the long-term depression pathway. As these last results have been partially reproduced in mice overexpressing wild-type or mutant TET1 catalytic domains in the hippocampus, TET1 appears to act as an “inhibitor” of long-term memory, independently of its catalytic function (Kaas *et al.*, 2013). While a consensus is hard to reach at this stage, it would seem that TETs affect both short- and long-term memory as well as NPC proliferation and differentiation. The discrepancies between Zhang’s and Rudenko’s results may be due to the use of different knockouts: Zhang *et al.* targeted exons 11–13 of the TET1 protein, whereas Rudenko *et al.* used knockout mice with a deleted exon 4 (previously described by Jaenisch and colleagues Rudenko *et al.*, 2013; Zhang *et al.*, 2013b). Further studies, including systematic conditional hippocampal knockout of TET1, TET2, and TET3 (and also double and triple knockouts), should help to assess these divergences and to distinguish the exact roles of the three TETs in neurogenesis.

Also beyond the neuronal context, reports show that the hydroxymethylcytosine pattern is highly dynamic during terminal cell differentiation, as described for preadipocytes, germ cells, and differentiating monocytes (Fujiki *et al.*, 2013; Gan *et al.*, 2013; Klug *et al.*, 2013; de la Rica *et al.*, 2013). During adipogenesis, the genome displays a global increase in 5hmC and local hydroxymethylation on genes that are later expressed in mature adipocytes (Fujiki *et al.*, 2013). A similar initial global increase in 5hmC, followed by a global decrease, and the presence of differentially hydroxymethylated regions (dhMRs) on certain coding and non-coding genes are also observed during spermatogenesis (Gan *et al.*, 2013). Finally, Klug *et al.* (2013) and de la Rica *et al.* (2013) have provided evidence that TETs, especially TET2, are essential to the proper differentiation of monocytes to osteoclasts: during this process, DNA demethylation preceded by 5hmC deposition is observed.

All in all, it thus seems that while 5hmC may decrease when mESCs lose their pluripotency, a global increase in this epigenetic mark can occur during terminal differentiation, sometimes followed by a subsequent decrease. This highlights the dynamicity of the hydroxymethylcytosine pattern and suggests a crucial role for the TET enzymes in cell differentiation and organ (e.g. brain) formation as well as a key role in cell reprogramming. With proteins so essential and so versatile, it comes as no surprise that their loss, or an alteration of the 5hmC pattern, can lead to diseases such as cancer or neurological disorders. Roles of altered TET functioning and hydroxymethylation in disease are discussed below.

**TETs/modified cytosines in cancer and non-cancer diseases**

TETs are on the one hand recognized as tumour suppressor and a hallmark of many cancers seems to be a significant decrease in 5hmC. This may be partly due to the well-known global hypomethylation that takes place during cell transformation, but in breast cancers, melanomas, and leukaemias without MLL rearrangements, reports have additionally revealed either a mutation in the TET2 gene (although the level of one TET does not always correlate with the global level of hydroxymethylcytosine) or substantial downregulation of all three TETs, for example by micro-RNAs (Tefferi *et al.*, 2009; Haffner *et al.*, 2011; Nestor *et al.*, 2011; Hsu *et al.*, 2012; Lian *et al.*, 2012; Yang *et al.*, 2012; Huang *et al.*, 2013; Song *et al.*, 2013a, b). In 2012, Hsu *et al.* (2012) reported TET1 to be downregulated in prostate and breast cancer tissues, and found its loss to facilitate cell invasion in xenograft mouse models, at least via deregulation of “inhibitor of metalloproteinase” proteins (TIMPs). In human melanomas, which show a marked genome-wide decrease in hydroxymethylcytosine and more subtle changes in methylation patterns, the most decreased TET is TET2. When TET2 is overexpressed in human melanoma cells xenografted onto immunodeficient NSG mice, tumour growth is suppressed. This suggests a crucial involvement of TET2 and 5hmC in melanoma development (Lian *et al.*, 2012). Furthermore, TET2 is often mutated in myeloid malignancies. In 4–13% of myeloproliferative neoplasms and 20–25% of myelodysplastic syndromes, somatic deletions and TET2-inactivating mutations are found, but do not correlate clearly with prognosis (Tefferi *et al.*, 2009; Bejar *et al.*, 2011). TET2 mutations are also found in 7–23% of AMLs, where they correlate with poor prognosis. A major role of TET2 in leukaemia has been confirmed in TET2-knockout mice, where loss of the enzyme appears to increase the haematopoietic stem cell compartment and to skew cell differentiation towards the myeloid compartment, causing symptoms resembling those associated with TET2 mutations (Li *et al.*, 2011; Moran-Crusio *et al.*, 2011; Quivoron *et al.*, 2011).

In fascinating contrast to the above, there are contexts where TETs appear to have an oncogenic action. They owe the designation “Ten-Eleven translocation enzyme” to a rare translocation of the histone methyl transferase gene MLL and of the TET1 coding region, giving the rare translocation partners (de Boer *et al.*, 2013). Yet Huang *et al.* revealed in 2013 that TET1 plays an essential oncogenic role in MLL-rearranged leukaemia. It is aberrantly overexpressed in MLL-rearranged AML, while the TET2 and TET3 expression levels remain unchanged. ChIP-qPCR and expression experiments have further shown that the TET1 promoter is a direct target of MLL-fusion proteins and that the increase in TET1 leads to a global genomic increase of hydroxymethylcytosine. Furthermore, MLL fusions and TET1 both co-regulate genes such as HOXA9 or MEIS1 that inhibit apoptosis and enhance cell proliferation, causing transformation and leukaemogenesis (Huang *et al.*, 2013).

In summary, the hydroxymethylcytosine level is often deregulated in cancer, but not always in the same manner: globally there may be an increase, as seen in MLL-rearranged leukaemia, or a marked decrease, as observed in many other cancer types. Recently, some elegant techniques have been created to precisely map 5hmC
at single-nucleotide resolution (Booth et al., 2012; Yu et al., 2012). This could make this epigenetic modification an interesting marker for use in diagnosing cancers, especially blood cancers, for which cell samples can easily be obtained.

Growing attention is also focusing on the involvement of TET dysregulation in non-cancer diseases. There is increasing evidence that epigenetic dysfunction and resulting changes in gene expression may contribute at least partly to the aetiology of various pathologies, especially neurological disorders (see Table 1) (Jakovcevski & Akbarian, 2012). As mentioned above, the mammalian brain is very rich in hydroxymethylcytosine. In Rett syndrome, mutations in MeCP2 impair the binding of this protein to 5mC, suggesting that the altered deposition of this mark by TET methyldioxygenases may also play a role in neurological diseases (Kriaucionis & Heintz, 2009; Mellen et al., 2012). Accordingly, several reports have revealed altered TET expression and an abnormal 5mC pattern in the brains of patients suffering from psychosis, Huntington’s disease, Friedreich’s ataxia, and fragile X syndrome (Guidotti et al., 2012; Yao et al., 2013). Psychosis is a broad category encompassing symptoms of schizophrenia and bipolar disorder. Psychotic patients have been found to display downregulated expression of several vulnerability genes, including GAD67, reelin, NR2A, and GAT1, in parallel with hypermethylation of the corresponding promoter regions (Dong et al., 2012). Recently, two groups additionally reported hyper-hydroxymethylation of the GAD67 promoter in the inferior parietal lobe, both in a psychotic cohort and in mice born with psychosis-like disorders (Dong et al., 2012; Mattrisciano et al., 2012). To evaluate the possible role of 5mC in GAD67 repression, it would be interesting to see if variation of the 5mC level in GAD67 might correlate with TET binding to one or more corepressors such as SIN3A, and NurD. In addition, Dong et al. (2012) revealed local variations in hydroxymethylcytosine content, a global increase in both TET1 and 5mC in schizophrenic and bipolar patients, and persons suffering from depression displayed a similar, albeit insignificant trend. A TET1 increase has likewise been observed in psychotic patients with a history of alcohol abuse and interestingly, in lymphocytes of schizophrenic patients. The latter finding suggests that it might be possible to identify peripheral methylation/hydroxymethylation biomarkers for use in diagnosis/prognosis when a biopsy is not an option (Guidotti et al., 2012; Auta et al., 2013).

Huntington’s disease (HD) is a genetic disorder characterized by expansion of glutamine residues in the protein huntingtin, due to accumulation of the CAG triplet in the first exon of the corresponding gene. The symptoms include difficulties in motor coordination and also psychiatric disturbances, cognitive disorders, and weight loss (Sadri-Vakili & Cha, 2006). In 2013, Villar-Menendez et al. (2013) showed that the promoter of the ADORA2A gene, encoding the G-protein coupled receptor A2AR known to be downregulated in HD and involved in the disease, displays a decreased 5mC level in a mouse model of HD and in the putamens of HD patients. In contrast to the global increase in hydroxymethylcytosine found in psychosis, dot-blot experiments and analysis of 5mC patterns by deep sequencing in an HD mouse model revealed a rather global decrease in 5mC, accompanied by a significant decrease in TET1 expression. Analysis of the differentially hydroxymethylated regions further uncovered an alteration of canonical pathways involved in neuronal development and differentiation pathways. Hence, the loss of hydroxymethylcytosine may impair neuronal function in HD brains and appears as a novel epigenetic marker of the disease (Wang et al., 2013).

Friedreich’s ataxia (FRDA) is another neurodegenerative disease due to expansion of a triplet (GAA). This expansion, here in the first intron of the FXN gene, leads to heterochromatin formation and gene silencing. Upstream from the repeats lay CpGs that appear hypermethylated in patients suffering from the disease. In the studies concerned, however, the techniques used to map 5mC could not distinguish this mark from 5mC (Campuzano et al., 1996; Saveliev et al., 2003; Al-Mahdawi et al., 2013). In 2013, Al-Mahdawi et al. (2013) addressed this issue, showing that nearly all the methylcytosine is in fact hydroxymethylcytosine in both FRDA and normal human cerebellar tissues, and that 5mC is more abundant in the former than in the latter. They further hypothesized that the observed 5mC increase might enhance production of FAST-1 antisense RNA, which could in turn mediate heterochromatin formation and FXN downregulation, causing the typical symptoms of the disease. Lastly, the fragile X genetic syndrome, triggered by FMR1 gene repression, is a widespread inherited cause of autism and mental retardation in boys (Tassone et al., 2000; Yao et al., 2013). While the disorder is known to be associated with hypermethylation of the FMR1 promoter and subsequent transcriptional gene silencing, possible global fragile-X-associated changes in DNA methylation/hydroxymethylation were not addressed until recently, when
Yao et al reported a global decrease in 5hmC in gene bodies and promoters and a more subtle increase in cerebellum-specific enhancers and some repetitive elements. The same report also presents dhMR analyses highlighting changes affecting functional pathways in neuronal development. As these results were obtained from a knockin mouse model, it will be important to assess the global level of hydroxymethylcytosine in fragile X syndrome patients (Yao et al, 2013).

In each of the above-mentioned neuronal diseases, the global level of hydroxymethylation thus seems to be altered, either upward or downward according to the disease. Also, some disorders may involve several tissues in the brain, and both the global and specific 5hmC patterns may depend on the analysed tissue. This is exemplified by two recent studies showing that in Alzheimer’s disease, the most common form of dementia in humans, hippocampal regions display a significant increase in 5hmC, while the middle frontal and temporal gyri show a significant decrease (Chouliaras et al, 2013; Coppieters et al, 2014).

In summary, the TET methylidioxygenase machinery seems altered in various diseases, notably cancers and neurodegenerative disorders but also probably other important diseases such as oxidative-stress related disorders (see modes of direct regulation of TET enzymes). This strongly suggests that 5hmC patterns are important in many pathologies. As DNA can be recovered easily from tissues and is stable over time, 5hmC profiles might be used to find clinically relevant epigenetic biomarkers. Yet questions remain to be answered. It will notably be important to identify clearly the pathways affecting 5hmC profile changes and those affected by them. We also need to distinguish whether variations of the global level of this mark are causes or consequences of the diseases concerned. These important questions will surely keep a lot of laboratories excited and busy for the next few years.

Concluding remarks

Since their discovery in 2009, the TET enzymes have sparked tremendous interest in the epigenetic community. Key studies have highlighted the importance of active or passive TET-primed demethylation in both cell reprogramming and cell differentiation, and the functional diversity of TET partner proteins suggests a more multifaceted action than anticipated. TETs associate with diverse chromatin-related machineries involved in transcriptional activation or repression, and the marks they create may act as epigenetic signals per se, i.e. as more than just intermediates in DNA demethylation. TETs also seem to play roles that do not depend on their catalytic activity.

Challenges abound in the world of TETs: untangling the intricate web of TET interactions with the chromatin environment, answering burning questions such as: is the frequently observed positive correlation of transcription with hydroxymethylcytosine in gene bodies the cause or the consequence of gene expression? What proteins act as true 5hmC, 5fC, and 5caC “readers” and what are their precise biological roles?

Figure 3. The evolving “TETris playground”.

A picture is emerging in which TETs build up specific “DNA modification blocks” (hmC, fC, caC) through their coordinated regulation at multiple levels. The following sequence of events can therefore be suggested: (i) Transcription factors will result in (ii) 5hmC, 5fC, and 5caC formation with 5hmC being more abundant in gene bodies. (iii) These modifications can be further bound by various “binders” or “readers” that will either participate in DNA demethylation or translate the epigenetic signal to transcriptional activation or repression. Other active players are transcriptional regulators and small molecules such as vitamin C that directly affect TET activity and hence the deposition of 5hmC, 5fC, and 5caC. These DNA modification blocks, TETs, and TET regulators can assemble in a gene- and/or a cell-state-dependent manner creating precise landscapes which will ultimately affect gene expression.
It comes as no surprise that perturbed TET functioning leads to disease, given the importance of TETs in so many key cellular processes. In many cancers a global loss of 5-hydroxymethylcytosine is observed, but it remains uncertain whether altered hydroxymethylation patterns, like altered DNA methylation patterns, can be viewed as true hallmarks of certain cancers. As various cancers display altered TET-gene expression or TET mutations, and as TETs play a role in generating 5c and c5c, the possible link between cancers and changes in 5c and/or c5c patterns is certainly worth investigating. This applies also to neuronal disorders, as several show a perturbed 5hmC distribution. There remain crucial gaps in our understanding of how TET dysfunctioning/dysregulation contributes to disease etiology, and filling them is likely to be rewarding. Once we do understand the link between TET functions and particular diseases, we can hope to develop therapeutic tools. Just as known DNMT-modulating drugs (e.g. selective non-nucleoside inhibitors) are commonly used to treat myelodysplastic syndromes (Sekeris & Cutler, 2013), TET modulators (inhibitors or activators) may prove clinically useful. Small molecules and TET-microRNAs antagonists could be promising candidates. In fact, the therapeutic use of microRNAs antagonists might be just around the corner: a phase II study is in progress on Miravirsen (Santaris Pharma), a miR-122 antisense Locked Nucleic Acid that significantly reduces the level of viral RNA in hepatitis-C-virus-infected patients (Janssen et al., 2013). Alternatively, TET enzymes might be targeted to specific genomic loci to restore the normal epigenetic landscape. Joung’s group has already exploited this idea, successfully using TALE-TET fusion proteins to target the TET1 catalytic domain to known methylated genes (Maeder et al., 2013).

No wonder the world of TET is in ferment! Although still under construction, a picture is emerging in which TETs both contribute to building “DNA modification block” patterns through creation and erasure of epigenetic marks and participate with diverse partners in a more intricate scheme. As results accumulate with their share of surprises, the simplistic view presented in Fig 3 is bound to become much more elaborate. The challenge ahead will be to fully relate the many levels in which TETs participate to the patterns they must properly create.

Author contributions
BD and FF wrote the manuscript, RD designed the figures.

Conflict of interest
The authors declare that they have no conflict of interest.

References
pluripotency and its loss is compatible with embryonic and postnatal development. Cell Stem Cell 9: 166–175


Lorsch RB, Moore J, Mathew S, Raimondi SC, Mukatira ST, Downing JR (2003) Tet1, a member of a novel protein family, is fused to MLL in acute myeloid leukemia containing the t(10;11)(q22;q23). Leukemia 17: 637 – 641


© 2014 The Authors

The EMBO Journal Vol 33 No 11 2014

1209


cooperatively regulate key genes for Xenopus eye and neural development. Cell 151: 1200–1213
Zhang P, Huang B, Xu X, Sessa WC (2013a) Ten-eleven translocation (Tet) and thymine DNA glycosylase (TDG), components of the demethylation pathway, are direct targets of miRNA-29a. Biochem Biophys Res Commun 437: 368–373